# AN ANALYSIS OF SCORING RUBRIC USE TO ASSESS THE STUDENT TRANSLATION AT D3 ENGLISH DEPARTMENT AT POLITEKNIK NEGERI MALANG

**THESIS**

**In Partial Fulfillment of the Requirement for**
**Master Degree of English Language Education**

**ANTON HARYADI**

**201710560211006**

**DIRECTORATE OF POSTGRADUATE PROGRAM**
**UNIVERSITY OF MUHAMMADIYAH MALANG**
**OCTOBER 2019**

**AN ANALYSIS OF SCORING RUBRIC USE TO ASSESS THE STUDENT TRANSLATION AT D3 ENGLISH DEPARTMENT AT POLITEKNIK NEGERI MALANG**

Proposed by:
**ANTON HARYADI**
**201710560211006**

Advisor I

**Dr. Estu Widodo**

Advisor II

**Dr. Sri Hartiningsih**

Director of Directorate of Postgraduate Program

**Akhsanul In'am, Ph.D**

Head of Program Study
Master of English Language Education

**Dr. Estu Widodo**

# THESIS

Written by

**ANTON HARYADI**
**201710560211006**

Has been examined in front of examiners
On **Tuesday, 29 October 2019** and decided that it
has fulfilled the requirements to get
Master Degree of English Language Education
In Postgraduate Program of University of Muhammadiyah Malang

**The Examiners**

| | |
|---|---|
| Chief | : **Dr. Estu Widodo.** |
| Secretary | : **Dr. Sri Hartiningsih.** |
| 1st Examiner | : **Dr. Sudiran.** |
| 2nd Examiner | : **Bayu Hendro Wicaksono, Ph.D** |

## LETTER OF STATEMENT

I, undersigned:

Name : **ANTON HARYADI**
NIM : **201710560211006**
Study Program : **Magister Pendidikan Bahasa Inggris**

hereby declares that:

1. The thesis entitled: **AN ANALYSIS OF SCORING RUBRIC USE TO ASSESS THE STUDENT TRANSLATION AT D3 ENGLISH DEPARTMENT AT POLITEKNIK NEGERI MALANG** is my original work and contains no one scientific paper that may be proposed to achieve an academic degree at my university. In addition, there is no other's idea or citation except those which have been quoted and mentioned in bibliography.
2. If this thesis is proven as form of **PLAGIARISM** in this thesis, I am willing to accept the consequences including accepting **THE CANCELATION OF THE GRANTING OF MASTER DEGREE** and undergoing any procedures required by the prevailing law.
3. This thesis can be used for literature review which can be accessed by others freely (**NON EXCLUSIVE ROYALTY**).

Thus, this statement is made truthfully to be used as appropriate.

Malang, 22 October 2019

The writer

METERAI TEMPEL
TGL 20
7E7F0AFF399953566
6000
ENAM RIBU RUPIAH

**ANTON HARYADI**

## MOTTO AND DEDICATION

**Motto:**

**There's always more than one way to do anything; just look for it.**

**Dedication:**

**Specially presented for:**

**My families**

**My lecturers**

**All my friends**

## ACKNOWLEDGEMENTS

Alhamdulillah, all praises to Allah for His blessing and the strengths in completing this thesis. I would like to express my sincere gratitude to all lecturers in University of Muhammadiyah Malang, who introduces me to a world of academic wonders beyond my expectation, especially for my advisors Dr. Estu Widodo, M.Hum and Dr Sri Hartininingsih, M.M for their patience, motivation, enthusiasm, and immense knowledge.

The sincere gratitude is also expressed to the Head of Department, Sugeng Hariyanto, and the faculty of D3 English Department at Politeknik Negeri Malang for the warm welcome and kind supports. I also want to thank my best friends "the Weekender" for the friendship, the joy, the laughter, and togetherness.

This work will also never be accomplished if the support was not given from the start by the Staff of TransKomunika, who provide a special permission for me to take the academic journey and finally complete it.

Last but not least, the same gratitude also goes to my wife (Dee) and kids (Kafka dan Keena), and also my parents for their endless love, support, and care during the tenure of this study.

Malang, 25 October 2019

Anton Haryadi

**An Analysis of Scoring Rubric Use to Assess the Student Translation at D3 English Department Program at Politeknik Negeri Malang**

**Master of English Language Education**
**Postgraduate Program University of Muhammadiyah Malang**
**antonharyadi@gmail.com**

## ABSTRACT

Assessment always plays a crucial role in any educational practices. In translation teaching, it is important as well in a sense that translation teaching requires a valid, reliable, and practical assessment instrument. In addition, it is also important to standardize every aspect to make sure a fair assessment for all. At D3 English Department Program at Politeknik Negeri Malang, the translator education is also organized to train the student translators. A uniform and agreed scoring rubric is required to make sure the assessment is fair and transparent. Therefore, this research investigated the current assessment practice and the expected scoring rubric to use in the future. It is expected that the problems can be identified and the solutions can be proposed.

In this research, descriptive qualitative research was selected to be the research design, with two research instruments, i.e. semi-structured interview guide and document analysis. This research was conducted at Politeknik Negeri Malang. Four research participants were involved in this study. The findings showed that the teaching and assessment practice in this vocational higher education institution has followed the ideal assessment process. However, the scoring rubric type and the frequency of use were still different from one lecturer to the other. In addition, the lecturers expected that the scoring rubric can be used for research. For the current scoring rubric, there were several shortcomings. Therefore, a new scoring rubric was required. The designing was expected to be using a top-down approach. An analytical scoring rubric was expected to be a new scoring rubric developed for them. They expected that the scoring rubric is more practical but still detailed and thorough.

**Keywords:** *Translation teaching, translation quality assessment, scoring rubric*

**Analisis Penggunaan Rubrik Pensekoran untuk Menilai Terjemahan Mahasiswa di Program D3 Bahasa Inggris Politeknik Negeri Malang**

**Magister Pendidikan Bahasa Inggris**
**Program Pascasarjana Universitas Muhammadiyah Malang**
**antonharyadi@gmail.com**

## ABSTRAK

Penilaian selalu berperan sangat penting dalam semua jenjang pendidikan. Dalam pengajaran terjemahan, penilaian juga tak kalah pentingnya karena pengajaran terjemahan juga memerlukan instrumen penilaian yang sahih, andal, dan praktis. Selain itu, setiap aspek pendidikan juga perlu distandarkan untuk memastikan semua pihak mendapatkan penilaian yang adil. Di Program D3 Bahasa Inggris Politeknik Negeri Malang, program pendidikan calon penerjemah dilakukan untuk mendidik para mahasiswa calon penerjemah. Rubrik penilaian yang seragam dan disetujui semua dosen diperlukan untuk memastikan penilaian adil dan transparan. Oleh karena itu, penilaian ini mendalami praktik penilaian yang berlangsung dan rubrik penilaian yang diharapkan untuk dapat digunakan di masa mendatang. Diharapkan masalah dapat diidentifikasi dan solusi dapat diajukan.

Dalam penelitian ini, rancangan penelitian yang dipilih adalah kualitatif deskriptif, dengan dua instrumen penelitian, yaitu panduan wawancara semi-terstuktur dan analisis dokumen. Penelitian ini dilakukan di Politeknik Negeri Malang. Empat partisipan penelitian dilibatkan dalam penelitian ini. Temuan penelitian menunjukkan bahwa praktik pengajaran dan penilaian di lembaga pendidikan tinggi berbasis vokasi ini sudah mengikuti praktik penilaian yang ideal. Namun, masih ada perbedaan antara satu dosen dengan dosen lain terkait jenis rubrik pensekoran dan frekuensi penggunaannya. Selain itu, para dosen berharap agar rubrik pensekoran juga dapat digunakan untuk penelitian. Untuk rubrik pensekoran saat ini, ada beberapa kelemahan terkait. Jadi, diperlukan rubrik pensekoran baru. Program perancangan skor pensekoran diharapkan menggunakan pendekatan *top-down*. Selain itu, mereka berharap agar rubric tersebut menggunakan pendekatan rubrik analitik. Mereka juga berharap agar rubriknya praktis tetapi tetap mendetail dan menyeluruh.

**Kata kunci:** *Pengajaran terjemahan, penilaian kualitas terjemahan, rubrik pensekoran*

## TABLE OF CONTENTS

## LIST OF TABLES

## LIST OF APPENDICES

## INTRODUCTION

### Background of the Research

Assessment always becomes an indispensable part, through which a teaching and learning process is bridged (Torrance & Pryor, 2001; William, 2013). This statement implies that there is always a gap between what the lecturers are teaching and what the students are really learning. Therefore, the lecturers' responsibility is to fill this gap so that the students can learn what they need to learn. As one of the tools to fill the gap, the assessment is aimed at bridging and minimizing this gap so a teaching activity really reflects the student needs.

In relation to this bridging function, Jabbarifar (2009) argues that assessment should include four basic functions: i.e. (a) measuring improvement over time, (b) motivating students to study, (c) evaluating the teaching methods, and (d) ranking the students' capabilities in relation to the whole group evaluation. The main purpose, as Jabbarifar puts it, is to measure the students' improvement over time, which should be effectively done in a regular basis. From the result of the measurement, the subsequent purpose is to give the students a feedback or benchmark, from which the students may be motivated. In addition, the assessment can be the basis to rank the students' capabilities for administrative purposes. Just like the students, the lecturers may also benefit from the assessment. The assessment result is an effective and a valid means to evaluate the effectiveness of the teaching methods employed by the lecturer at the time.

This opinion is in line with the assessment purpose categorization, i.e. diagnostic, formative, summative (Gavotto-Nogales, Morales, & Pierra, 2015; Melis & Albir, 2001) and evaluative (Black & Wiliam, 2018). Diagnostic assessment refers to a range of assessment processes to diagnose strengths and needs of students in a given learning field. Hence, it takes place before the learning process is started. Formative assessment refers to a variety of assessment methods to conduct during the teaching and learning process, while summative assessment is an assessment process after the teaching and learning process. In addition to the three assessment types, the fourth assessment is evaluative assessment. As the name implies, the purpose is to evaluate the whole teaching and learning process to make sure that it can be improved well.

Just like other field of skill based teaching, the assessment also plays an important role in translation teaching. The assessment is a helpful aid not only for the lecturer to reflect on their teaching, method, material, and assessment but also for the students to monitor their own progress and motivate themselves. In this respect, it is in line with assessment types according to its functions. Just like speaking and writing in terms of its nature as productive skills, the translation task and test are subjective for authenticity and validity purposes. In an attempt to minimize the subjective bias and move toward objective assessment, the scoring rubric is favored (Angelelli, 2009; Garant, 2009; Khanmohammad & Osanloo, 2009).

The scoring rubric itself can be grouped into two main categories, i.e. analytical and holistic scoring (Brookhart, 2013). Richards & Schmidt (2002) define analytical scoring as a scoring method that separates and weights different features of the test taker's performance and then assigns separate scores to each feature. It means each component of translation is assessed and weighed in a separate manner. On the contrary, holistic scoring refers to a scoring method in which a single score is assigned to the task samples on the basis of an overall impressionistic assessment of the test taker's performance. In this view, the translation is considered a whole discourse, which should not be assessed in a separate manner.

As a new department founded on 2017 under Politeknik Negeri Malang, D3 English Department has not used any single agreed scoring rubric. This is because the department is new, and most lecturers graduated from English language teaching rather than translation study program and have a relatively few translation experience. Thus, they need to adjust themselves to their new tasks in teaching translation. In this regard, the department can benefit from the previous research. Although, as argued by Melis & Albir (2001) and (Angelelli & Jacobson, 2009), the number of empirical studies is a relatively few in relation to assessment within university translation programs. From the relative few studies, there are several previous researches relevant to this research.

First, Hariyanto (2009) develops a scoring rubric in a section of his dissertation. He developed a scoring rubric using top down approach by adopting the integrated translation model developed by Bolanos. The rubric itself was used to evaluate the translation products found in several websites in Indonesia. However, this rubric is designed for a research purpose rather than a pedagogical purpose. Hence, it lacks the practicality aspect, which is really needed in pedagogical field.

Second, Garant (2009) employed a linear study in English translation program of University of Helsinki in 1997, 2001, and 2008. He found that point based error grading has now been replaced by a holistic grading method in this institution. He argues that the holistic grading tends to be viewed as more relevant to training future translators for real world tasks. This study, however, was carried out in a European city where English is not positioned as a foreign language. Therefore, the result of this study may not be applicable in Indonesian university contexts.

Third, Khanmohammad & Osanloo (2009) studies about the translation lecturers' opinion about which translation components should be incorporated in the rubric. There are 36 out of 50 participants taking part in this research. They found out that the most of the participants advocate the objective assessment of student translation using the scoring rubric. In addition, they provide their feedback about which translation components should be included in the assessment rubric. This research was situated in an Asian city which shares a lot in common with Indonesian universities. However, the research does not investigate other components in the scoring rubric, such as translation component/error description, error weighting, error severity categorization, etc.

Based on the elaboration above, this research is intended to fill these gaps; i.e. (a) it is to know how the current assessment practice is taking place, and (b) it is to contribute to the relatively few empirical studies on the assessment for translation program at a university level.

**Statements of the Problems**

Informed by the rationale above, the research problems are formulated below:

1) How is the assessment taking place at D3 English Department at Politeknik Negeri Malang?
2) How is a scoring rubric designed to assess the student translation at D3 English Department at Politeknik Negeri Malang?

## REVIEW OF RELATED LITERATURE

### Procedures of Designing a Scoring Rubric

For many educators, a scoring rubric is beneficial to assess the students' performance in certain fields of study. Therefore, they propose their own version of how a scoring rubric should be designed. Their proposed design procedure can be roughly categorized into two main approaches (Brookhart & Nitko, 2015). Although the two approaches do not necessarily lead to the same rubrics, they can be adjusted in accordance with purposes. The further explanation is elaborated below.

### A. Top-down approach

As the name suggests, a top-down approach is deductive in nature. As Brookhart & Nitko (2015) suggests, a conceptual framework becomes a starting point and describes the content and performance to be assessed. A conceptual framework here usually refers to the standard or theories which become the basis of scoring rubric design and development. It is best exemplified in Hariyanto's (2009) scoring rubric, in which the Dynamic Translation Model (DTM) bases his rubric development.

Below are the steps in the top-down approach, as suggested by Brookhart & Nitko (2015):

1. Create (or adapt from an existing source) a conceptual framework for achievement. This conceptual framework should describe the intended achievement (in this case for example, what is a good translation?) and outline the qualities that the students should learn and be asked to demonstrate (the achievement dimensions). It is advisable that the outline describes the continuum of performance in each criterion.
2. Write a general scoring rubric using these dimensions and performance levels. In this step, the scoring rubric designer should decide whether the scoring rubric will be designed analytically (one scale for each criterion) or holistically (one scale considering all criteria simultaneously).
3. Write performance descriptions at each level. After the general rubrics are completed, the performance should be described clearly and comprehensively to make sure the scoring rubric is easily understood.

4. Adapt the general scoring rubrics when it is needed for a specific learning objective. For example, if the rubrics say "The student shows understanding in linguistic principles," the specific rubric may be revised to say "The student shows understanding in selecting the accurate spelling variation."
5. Use the generated rubrics to assess several students' performances. After multiple- use, the feedback can be gathered from the students, which can be used to revise them for final use.

In short, those five steps are necessary in designing a scoring rubric based on the existing theory. However, the steps may change in accordance to the situation in the field.

**B. Bottom-up approach**

The second approach is a bottom-up approach, which as the name suggest is inductive in nature. Brookhart & Nitko (2015) suggests that this approach can be started with samples of students' work and uses them to create a framework for assessment. In addition, they say that this approach can be used when students are involved in designing their own assessment. However, this approach is not applicable in an Indonesian undergraduate university context, where the students are often not aware of the performance expected from them.

This bottom-up approach however can be used to design the rubrics which based on the translation lecturers' feedback on the translation quality. This was used in Khanmohammad & Osanloo's (2009) rubric, in which they collected the opinions from the translation lecturer throughout Iranian universities. Below are the steps:

1) Collect the translation lecturers' opinion about important components in assessing the translation. The components should be relevant to the kind of performance expected from the students. If possible, the component should be intended to assess the general learning outcomes.
2) Sort the components based on the most important, important, and less important. It is to make sure the most important components are put in the first rank, the second in the second rank, etc.
3) Assign the highest point or weight to the most important translation components suggested by the lecturers.

4) Write specific descriptions of why each piece of work is categorized as it is. The description should be specific in order that the rubric can be used by any lecturers interested in using the rubric.
5) The last, use the generated rubrics to assess several students' translation. After multiple try-outs, the feedback can be collected from the students to make it ready for final use.

In short, although it is not always the case, these five steps need to be taken to design a scoring rubric based on the lecturer's feedbacks. However, this approach may not be appropriate when the lecturers are relatively inexperienced in translation and translation teaching.

**Translation Quality Assessment**

Assessment plays an important role in any educational program, including translation teaching. In translation field, this assessment is widely known as translation quality assessment or TQA (Benhaddou, 1991; House, 2015; Medadian & Mahabadi, 2015). This sub-field of translation studies is not only growing fast (Barghout, 1990; Ghafouripour, Branch, Eslamieh, & Branch, 2018; House, 2015), but also delicate (Zehnalová, 2013).

House (2015) defines it as an activity "both retrospectively assessing the worth of a translation and prospectively ensuring the quality in the production of a translation." Based on this definition, the assessment starts by assessing whether the translation is worthy or not. After assessing its worthiness, the subsequent action is to ensure the quality in the production so that the final translation is worth publishing. This definition seems to be appropriate for industrial rather for pedagogical purposes as the student translation is not intended to be readily publishable. If the translation quality assessment lacks pedagogy knowledge, the result will be a skewed and subjective assessment (Shiyab, 2013).

In this regard, Hariyanto (2009) recommends that the translation quality assessment should divided into three parts according to its purpose, i.e. TQA for research, industry, and pedagogy. It is based on the belief that each starts from different perspectives and serves different purposes. TQA for research purposes are constructed using a certain theoretical frameworks which inform the whole research undertaking. TQA for industry usually starts from general linguistics and error typology, while TQA for pedagogy is similar to industry. In this research, this division is adopted. Therefore, this review of related literature focuses on TQA for pedagogy. Hariyanto (2009) suggest that there are at least four characteristics of pedagogical TQA. They are (1) may be based on theory but not very closely, (2) error identification, (3) error categorization, and (4) pattern of errors. This TQA was in form of a scoring rubric.

**A. Definition of Scoring Rubric**

Feedback is the heart of the assessment either for lecturers or students. It provides a significant input for the lecturers and students to know which areas of constituent skills they need to focus on. To provide meaningful and purposeful feedback, it is expected that the data based on which the feedback is given should be objective and transparent. The objective and transparent data help the students trust the feedback. To shift from subjective to more objective perspective, scoring rubric is deemed as one of the efficient solutions.

Scholars mostly think that scoring rubric is beneficial, and generally define rubric definition in an almost similar manner. Stevens & Lewvi (2005) write that a rubric at its most basic is a scoring tool, which lays out the specific expectations for an assignment. They emphasize that the rubric should clearly define the expectations set for a specific assignment. The more complete definition is given by (Popham, 1997) that the rubric refers to any assessment tool which lists criteria and provides some explanation for levels or gradations of quality. In his definition, some explanations should be given in the rubric to complement the criteria or the expectation.

The most comprehensive definition is given by Richards & Schmidt (2002). In the dictionary they both designed, a scoring rubric is defined as a set of scoring criteria or guidelines which are used in scoring or judging a test taker’s product, performance or response to a constructed-response item type of assessment task to make raters’ subjective judgments more reliable. The additional emphasis is put on the purpose of rubric, i.e. to make raters’ subjective judgments more reliable.

Based on the three definitions given above, it is agreed that the rubric should consist of criteria or expectation, explanation, and purpose. Thus, these three components should always be embedded in the scoring rubric since it is the center of the scoring rubric itself.

## B. Types of Scoring Rubric

Just like the other productive skills such as writing and speaking in English language teaching, translation can only be tested using a subjective test format for validity purposes. To make it more objective, scoring rubric is highly required to make sure bias and subjectivity does not impair the score and feedback. Before further discussing the scoring rubric, it is important to know the types of the scoring rubric.

First is analytic scoring rubric. Richard and Schmidt (2010) defines analytical scoring as a method of scoring which separates as well as weights different features of the test taker's performance on a specific task and assigns separate scores to each feature. It is clear from this definition that analytic scoring rubric should consist of separate features of competence. The common features in writing, for example, are content, organization, cohesion, style, register, vocabulary, grammar, spelling, and mechanics. The scores are assigned separately to each of these features, which will make up the total score.

Second is holistic scoring rubric. Richard and Schmidt (2010) defines holistic scoring as a method of scoring in which a single score is assigned to the samples on the basis of an overall impressionistic assessment of the test taker's performance. In contrast with analytic scoring where each feature of the product or performance is scored separately, only a single score is assigned to a whole sample in a holistic scoring. Thus, there are explanation and description which can be picked by the rater and its corresponding score.

Both scoring rubrics are the most common scoring rubric types, with which the lecturer scores the student translation. It is up to the lecturers who can decide the scoring rubric type they are going to use. It is important however to review the advantages and disadvantages of each type.

## C. Advantages and Disadvantages of Scoring Rubric

The scoring rubric has been widely used in Indonesian education lately. The curriculum designer and the policy maker seem to be aware of the advantages of the rubric for the education. However, some lecturers just take it for granted. One possible reason is that the lecturer is not aware about the advantages of scoring rubric. Therefore, it is important to lay out the most common advantages as well the disadvantages of scoring rubric.

Stevens & Lewvi (2005) elaborates some advantages of scoring rubric. First is that scoring rubrics provide timely feedback, prepare students to use detailed feedback, encourage critical thinking, facilitate communication with others, help the lecturer to refine the teaching skills, and level the playing field. The rubric benefits not only the teaching and learning process, but also the students and even the lecturer. The rubric even creates an interactive opportunity among students, between students and lecturers, and even between the task and the students. The interaction is expected to create a meaningful and purposeful learning experience for the students.

Despite the advantages, the rubric also has some weaknesses. First is about the students' free exploration in their learning. The rubric limits the students only to what is stated in the rubric. They will not take initiative to creatively solve problems. Second is that the rubric may demoralize the students if they find the rubrics too complex. The third is that the rubric design and development may be time consuming for the lecturers. In short, the rubric should be designed in that it minimizes the negative effects stated here.

In short, rubrics are proved to have more advantages over the disadvantages. The lecturer and rubric developer should be able to design an accurate rubric which minimizes the disadvantages. Therefore, the rubric is worth trying to enrich the learning experience.

**D. Examples of Scoring Rubric**

Assessing the translation students' works is never an easy task for the lecturer. Some lecturers even find it difficult to differentiate among translation evaluation, criticism, and assessment. To facilitate easier assessment, some lecturers use holistic assessment, while some lecturers use analytic scoring. Although there are controversies around the benefits of using holistic and analytic scoring, some lecturers prefer analytic scoring to let the students know the areas for improvement. To reduce the subjectivity in the scoring process, some lecturers adopt and adapt the scoring rubric. In practice, there are several scoring rubrics.

First, a translation scoring rubric is proposed by Hariyanto (2009), using the Dynamic Translation Model developed by Bolanos. Bolanos himself describe translation as bilingual communicative process which includes ST writer textualization, translator's comprehension and textualization, and last is the TT reader comprehension. This model bridges the gap between the ST oriented approach and TT oriented approach. There are four criteria used in scoring the translation, i.e. pragmatic, semantic, stylistic, and semiotic dimensions.

Second, it is a model of translation quality assessment (TQA) developed based on Waddington (2003). In Waddington's TQA model, there are four methods of assessment. The first method is more popular than the other three. There are three categories under the first method, they are (i) Inappropriate renderings which affect the understanding of the source text and are divided into eight categories: counter-sense, faux sens, nonsense, addition, omission, unresolved extra-linguistic references, loss of meaning, and inappropriate linguistic variation (register, style, dialect, etc.), (ii) Inappropriate renderings which affect expression in the target language and are divided into five categories: spelling, grammar, lexical items, text, and style, and (iii) Inadequate renderings which affect the transmission of either the main function or secondary functions of the source text.

The third translation rubric is proposed by American Translators Association (ATA). It is called ATA Framework in a form of table, which lists 22 criteria for errors. Developed mainly for ATA certification purpose, this framework has also been used by several lecturers across the globe in the translation classroom, for example Cuc (2017). The list is quite extensive, but it can only be used when the text length average more than 250 words. This framework, however, should be adapted when it needs to be implemented in Indonesia. One of its criteria, such as diacritical mark/accent, is not applicable in Indonesia.

All in all, these three rubrics discussed here indicate that translation quality assessment has been studied extensively due to its importance in the translation education, translation quality assurance, and even the translator certification. Therefore, it is necessary for the education institution implement a scoring rubric in their assessment practices.

## RESEARCH METHOD

### Research Design

In an attempt to answer the research question, a research plan should be devised in such a way that it agrees with the nature of problem being scrutinized. In this light, the plan is called research design (Ary, Jacobs, Razavieh, & Sorensen, 2010; Leavy, 2017). Ary et al. (2010 p. 22) say that an educational research can be broadly classified into two major categories: quantitative and qualitative research.

Informed by the purpose of the study, qualitative approach was employed in this research. This choice is informed by several reasons. First, a classroom assessment was a complex process, of which the solution requires a deep analysis. Second, the data were mostly in form of words, existing documents, and questionnaire. Third, the research unfolded in line with the progress of the research.

This research sought to portray human actions, wish, opinions, and etc. in their own context to understand the complex patterns in a sufficient detail and depth. Based on the portrayal, sufficient data were collected to understand the problem being investigated. In this sense, the problem was to understand the current assessment practice. Upon understanding the problem, an investigation was focused on how the lecturers design the scoring rubric to minimize the subjectivity. Therefore, descriptive qualitative was employed in this research.

### Research Setting and Participants

As elaborated earlier, the study was inspired by the absence of a single agreed translation scoring rubric which meets the criteria of empirically tested, practical, in accordance with the lecturer needs, and follows the industry practice. This research was situated at Politeknik Negeri Malang, especially at newly founded department, the English Department. The four translation lecturers were selected as the research participants.

The first participant was Mr. SGH, holding a doctorate degree in English language teaching with a dissertation on translation. He has been teaching English for more than 10 years, and translation subject for more than 5 years. Currently, he was teaching “General and literary text translation.” In addition to being an academician, he was also a seasoned translator with more than 10 years. The second participant was Mrs. HC, a doctorate degree in applied linguistics. She has been teaching English at university level for more 10 years, and translation subject for more than 1 year. Currently, she was teaching Subtitling. Her translation experience spanned for more than 5 years. The third participant is Mrs. FK, holding a master degree in English language teaching. She has been teaching English at university level for more than 5 years, and has been teaching translation subject for more than 1 year. Currently, she was teaching Introduction to Translation, General and Literary Text Translation, IT for Translation, Subtitling. The fourth participant is Mrs. RRT, holding a master degree in English language teaching. She has been teaching English at university level for more than 10 years, and translation subject for more than 1 year. Currently, she taught Subtitling.

**Research Instruments**

To collect the data from the lecturer, there were at least two instruments required for this research.

**A. Interview Guide**

In this research, in-depth interview was used for several reasons; (1) it allowed detailed data collection about person's thoughts and behavior, (2) it could give context to other data, and (3) it enabled multiple data collection from multiple interview session. This was to gather data on the lecturers' opinion on the current assessment practice, a scoring rubric use, and the purpose of using scoring rubric. In addition, the interview was to gather data on the frequency of assessment, and what steps they currently took to assess the student translation. The last, the interview was to gather data on the aspects of translation or assessment characteristics they considered most important hence should be prioritized.

The interview was semi-structured and open ended to make sure the lecturers' opinion and thoughts were shared freely. The interview guide was in Appendix I. The interview was the main data in this research since it captured the research participant opinion.

## B. Document Analysis

In this research, document analysis was used to complement the data. This instrument was used due to several reasons as suggested by Bowen (2009). First, documents are practical and manageable resources. Second, it is non-reactive data source, which can support the interview data. Third, it is beneficial to contextualize and support the research. The data were used to triangulate the data from the interview.

There were three documents in a form of scoring rubrics employed by the lecturers at Politeknik Negeri Malang. The documents were matched against the interview results to confirm if the data were credible.

## Data Collection Procedure

This research employed two research instruments to collect the data. The instruments were then used to collect the data based on the following procedure.

1 Selecting the research participants based on the proposed criteria.

2 Conducting interview session with the translation lecturers at D3 English Department at Politeknik Negeri Malang about how the assessment takes place and how they want to design a new scoring rubric.

3 Conducting a document analysis. After the interview session was completed, the document analysis was carried out to confirm if the lecturer opinions and statements were in line with the documents currently on file.

## Data Analysis

Upon the data collection, the next step was to analyze the data. In this research, the thematic analysis framework developed by (Braun & Clarke, 2006) was used to analyze the data. There were at least six detailed steps in data analysis. In this study, the data from the qualitative interview and document analysis were almost the same. To make sure the data can be cross-checked, they were analyzed in a single data analysis cycle.

To adjust the data analysis to the current research, the thematic analysis steps were modified below.

**Table 1: Thematic Analysis Steps**

| No. | Phase | Description of the process |
|---|---|---|
| 1. | Familiarizing the researcher's self with the data: | • Transcribing data recorded during the interview.<br>• Reading and re-reading the transcribed data and the existing documents to make sure its intactness and completeness. In addition, the contextual note was supplemented to emphasize its contextual relevance to the whole research.<br>• Noting down the initial ideas generated during the reading the re-reading phase in a separate book to gain the general picture of the research results. |
| 2. | Generating initial codes: | • Coding interesting features found in the data systematically across the entire data set. The coding used an underlining technique, where each feature was underlined.<br>• Collating the relevant interesting features of data to each different code. |
| 3. | Searching for themes: | • Collating the codes into several potential themes. The coded data were checked to make sure it belonged to the potential themes.<br>• Gathering all data relevant to each potential theme. To support if the theme was consistent, the relevant data were identified and gathered to backup the theme. |
| 4. | Reviewing themes: | • Checking if the emergent themes were in line with the coded extracts (Level 1) and the entire data set (Level 2).<br>• Generating 'map' of the analysis thematically. |
| 5. | Defining and naming themes: | • Refining the specifics of each theme continuously in relation to the overall story.<br>• Giving names for each theme and producing clear |

| | | |
|---|---|---|
| | | definitions. |
| 6. | Producing the report: | • Selecting clear extract examples from each theme.<br>• Relating the result of the analysis to the research questions and literature review. |

**Triangulation**

To avoid bias and make sure that the research is valid and reliable, there are two trustworthiness strategies used in this research:

**A. Credibility**

To meet a credibility criterion, member check was used as the strategy to improve the accuracy and eliminate the misunderstanding. When the data collection complete, the researcher confirmed several unclear phrases or words with the lecturer. To identify inaccuracies, help collect additional data and avoid misunderstanding, the unclear interpretation of the data was also be shared with the lecturer.

**B. Confirmability**

As the popular strategy to make sure confirmability criteria is met, audit trail was also used in this research. The audit trail is undertaken by highlighting the unique and interesting topics during the transcription, writing down the comments about coding, provide a rationale for merging the code, and explaining the meaning of the themes.

## FINDING AND DISCUSSION

### Research Finding

This section presents findings in relation to how the assessment practice was taking place. The discussion on how the scoring rubric was designed follows.

**A. Current Assessment Practices**

As discussed earlier, translation training is more toward skill-based training rather than knowledge-based training in nature. Although the general translation theories are still taught, the exercises to improve the student translation skills should be prioritized and multiplied. Therefore, the teaching in the institution was attempted to improve the student translation skills.

First, it was found out that all four lecturers preferred giving students a lot of translation exercises/assignments based on the data from the interview. One lecturer further explained that “I assigned a lot of exercises, so the students practice translating texts most of the time. (Appendix 2, SGH, Line 4-5).” Based on this interview, it indicates that the lecturer gave a lot of translation exercises to the students so that the students practice and improve their translation skill. It also means that most of the class time was spent on translation exercise rather than other activities.

This lecturer’s opinion was supported by another lecturer as she said that “It’s translating texts, based on the module (Appendix 2, RRT, Line 84).” Based on this interview excerpt, it was clear that another lecturer was also similar in which she also assigned translation exercises to the students. It implies that the lecturers understood the importance of skill over the knowledge.

Second, the next stage, upon the submission of the exercise, was to correct or assess the exercise papers/results. Based on the interview, it showed that the lecturers used a mixed use of assessment/scoring technique. One lecturer explained that “I use a scoring rubric, a combination of both, analytical and holistic” (Appendix 2, FTR, Line 45). This interview excerpt indicates that this lecturer used two scoring rubrics, i.e. a combination of analytical and holistic scoring rubric. It means that the lecturer already know that the subjective task such as translating text should be followed by an objective means of measurement, namely the use of a scoring rubric. To make sure the maximum benefit for the students, she used a combination of analytical and holistic.

This opinion was supported by another lecturer. He states that “For the quiz, more intuitively for the quiz. But for the mid test there is a rubric” (Appendix 2, SGH, 176-177). From this information, it is clear that this lecturer also made use of a scoring rubric although it came in different types. Later, he provided an example of an analytical scoring rubric, which can be found in Appendix III. The ATA Framework rubric is originally designed by the American Translators Association for certification purpose there. In an unofficial unrecorded circumstance, the lecturer said that this was proven to be beneficial for translator training purpose.

Third, the frequency of the scoring rubric use was also different from one lecturer to the others. In regard to when the lecturer employed a scoring rubric, their answers can be categorized into three: only in summative test, in formative test, and every exercise the students completed. One lecturer explained the scoring rubric was used when "this is in the final and the middle test... (Appendix 3, SGH, 162)." This implies that this lecturer used a scoring rubric only in summative test. In the other part of interview, he told that every meeting usually involves the class discussion on the student translation. Therefore, it was not necessary for the lecturers to use a scoring rubric in every meeting.

However, the other lecturer gave a different response, in which she said that "every time students have exercise. (Appendix 2, FTR, 51)." Thus, it can be said that this lecturer was diligent in a way that she employed a scoring rubric every time the students completed their exercise. If there was one exercise every week, the lecturer would correct the same amount every week. It may be the reason why the lecturer really needed a practical scoring rubric.

Fourth, it is important to know the reasons why the respective assessment instruments were chosen by the lecturers. They agreed that it is for due practicality reason. One of the lecturers explained that she used a scoring rubric because it is "Easier to score and it is less subjective, allowing me to use the results for a research article as well later. (Appendix 2, RRT, Line 89-90)." Based on the excerpt above, it can be concluded that the use of a scoring rubric was not only to facilitate an easier scoring, but also to reduce the lecturer's subjectivity or bias toward certain student translations. One interesting highlight is about the possibility of using the scoring data for research article.

In relation to this, one seasoned lecturer who happened to be a translator scholar further added that "the research in that case is the research on the teaching of translation. Not the research on the translation itself, but the teaching of the translation. (Appendix 3, SGH, Line 181-183)." It means that it is possible that the data were used for research, but only limited to the research on translation teaching. It is quite new since the assessment is typically categorized into four purposes, i.e.

diagnostic, formative, summative, and evaluative. Thus, this new purpose can be categorized as new finding.

Fifth, all the lecturers thought that they needed to share the assessment criteria with the students. In regards with the reasons why they shared it with the students, they had almost the same answers. They want the students to have a fair assessment on their skill/understanding on what they learn. One lecturer said that this is:

> To inform the students about which part of their translation which needs to be improved and to let them know how I do the scoring/grading. Also, to let them know which parts of their translation are evaluated. (FTR, Line 55-57).

Based on the opinion above, it is clear that the students needed to know about which aspects of the skills they needed to improve and how she did the scoring/grading. This way, the students would know from the first what were expected from them. With this knowledge, they could perform in line with the lecturer expectation or even outperform it. In addition, it is important for the students to feel a sense of fair assessment, which may directly or indirectly contribute to the students' motivation. Another lecturer supported the above statement. He further explained that a scoring rubric was used in order "To make it fair and so that the students will be aware of their own performance during the test and thus perform their best (RRT, 98-99)." In short, they reached an absolute agreement that the students needed to know the assessment criteria in the instruments.

Sixth, there were several weaknesses currently found by the lecturer at Politeknik Negeri Malang in the rubric. The weakness here is related to the scoring rubric itself and the designing of scoring rubric. In relation the rubric, one lecturer explained that "sometimes they are not quite thorough because I adapt other people's rubric. (Appendix 2, FTR, Line 66)." Based on this lecturer's statement, the obstacle lied in the scoring rubric, which was not quite thorough. The last obstacle was stated by another lecturer, who said that that the scoring rubric development "takes time to make and prepare. (Appendix 2, HLD, Line 128)." Based on this excerpt, it is clear that the process of designing and preparing the scoring rubric takes time. Based on the interview results, the thoroughness here referred to at least five issues, i.e. (1) not straightforward, (2) not precise, (3) not quite thorough, (4) not standardized, and (5) time it took to make and prepare the good scoring rubric.

In short, the assessment procedure and mechanisms were different from one lecturer to the other. It might be due to the lecturer different teaching style, education background, seniority, and courses they were teaching. Despite the differences, practicality consideration or ease of doing were responsible for their choice of assessment practice. Practicality in this context is not only embedded in the rubric, but also in the assessment practice.

**B. Designing of a Scoring Rubric**

A scoring rubric plays a vital role in translation teaching, especially in an attempt to improve the student translation competence. It serves as a feedback to the students and the lecturer, as a score for a formative purpose, and an input to the faculty to later improve the curriculum. In this institution, the same awareness was gained by the lecturers, which thus attempted to have a uniform, agreed and standardized scoring rubric. In an attempt to do that, several opinions were given by the lecturer.

First, all the lecturers agreed that the scoring rubric should be based on the existing analytical scoring rubric. It can be seen from one of the lecturer statement that "I think we should adapt the existing scoring rubric. (Appendix 2, HLD, Line 135)." This excerpt indicates that the lecturer preferred adapting the existing scoring rubric. It seems that the head of the institution has not decided what types of scoring rubric to use and the lecturers have not agreed on a single scoring rubric. This was confirmed by the following lecturer as well as the chief of the department.

> We should but we did not do that yet, I think every lecturer making his or her own way for the practicality and it's not uniform yet. We do not have any single agreed scoring rubric to score the students translation (SGH, Line 341-343).

This opinion suggests that a single scoring rubric was not still agreed by all the lecturers. They still employed a different rubric depending on their own need. However, that they had the same opinion about the choice of approach in designing a scoring rubric. Among the three choices, such as develop, adapt, and adopt, all the lecturers prefer an adapting approach to the other two approaches. It means that the lecturers need to select an appropriate and standardized scoring rubric, on which the adaptation need to be based. The selected scoring rubric is then adapted to create a single standardized scoring rubric to use in the institution.

Second, all the lecturers had the same opinion about which types of scoring rubric should be designed and developed, and later be employed by the lecturers. One of the lecturers said that "I think it is a rubric, where each translation component is scored separately. (Appendix 2, FTR, Line 72)." It can simply be stated that an analytical scoring rubric was preferred. Another lecturer supported this statement, where she said that "I will pick analytical scoring rubric. (Appendix 2, RRT, Line 116)." Based on these two excerpts, an analytical scoring rubric was chosen by the lecturer if the new scoring rubric should be adapted. It was consistent with their previous answers that the scoring rubric they expected was a detailed and a thorough one. The "detailed" nature is identical with an analytical scoring rubric, in which the rubric should detail each component and criteria being assessed while the holistic is more general.

Third, they had different opinion about which characteristics to prioritize in designing a scoring rubric. The following table shows their different statements:

**Table 2. Priority of the New Scoring Rubric**

| No. | The Teacher Answers | Line |
|---|---|---|
| 1. | Practicality | SGH, Line 233 |
| 2. | Practicality | FTR, Line 77 |
| 3. | Practicality | RRT, Line 121 |
| 4. | Reliability | HLD, Line 138 |

Based on the table, it was clear that three lecturers chose practicality, and one lecturer chose reliability. Based on the table, it can be said that practicality became the first priority, and reliability became a second priority in terms of the lecturer preference. To confirm if this was the case, the second interview was administered to the head of the department. He stated that:

> I think because if it is too detailed, it is really discouraging because we have to read then we have to mark and we have to convert it into score and we have a lot of students… (Appendix 3, SGH, Line 331-333).

Based on this interview, it can be deduced that a too-detailed scoring rubric would discourage the lecturers because there were still a lot of accompanying activities, i.e. read the translation, mark it, and then convert the results into the scoring rubric. Unfortunately, the size of the students is also large. Therefore, it is important for the lecturers to have a practical scoring rubric.

Fourth, there were several suggestions made by the lecturers for future scoring rubric. The suggestions were related to the scoring rubric itself and could be categorized into four sections, i.e. more detailed, thorough criteria, appropriate and proportional weighing, and meeting the needs of the students. In relation to this, one lecturer stated that the future scoring rubric should be "More detailed and thorough criteria should be provided and they should meet the needs of the students. (Appendix 2, FTR, Line 68-69)." Based on this excerpt, it is clear that the improvement was still expected from the existing scoring rubric the lecturer employed now. The improvement regarding the detail and the thoroughness level was still expected and it also needed to be student-centered. This is important to make sure the scoring rubric is always improved.

This suggestion was strengthened by another lecturer when said that "It should be more detailed and bear appropriate and proportional weighing on its different components. (Appendix 2, RRT, Line 112-113)." Based on these two opinions, it can be stated that the lecturers just expected a more detailed scoring rubric. Their answers are consistent with the previous answers that the existing scoring rubric did not meet their expectation. However, the term "detailed" itself is too broad, which therefore should be described further.

All the three lecturers had some interesting points about the term "detailed", i.e. (1) a thorough criteria, (2) clear descriptions on which competence/sub-competence are evaluated, (3) clear and unbiased score ranges, and (4) appropriate and proportional weighing. In short, the detailed scoring rubric was really required by the lecturers to make sure their assessment was appropriate.

In short, there are four factors to consider while designing or adapting the analytical scoring rubric in the future. Since the approach the lecturers chose was adapting the scoring rubric, they should be consulted before a new scoring rubric is designed. It is important to make sure that the resulting scoring rubric is practical and relevant to the educational field. Thus, the new scoring rubric is more likely to be accepted and used by the lecturers.

## Discussion

This section presents a review the findings against the previous research findings and the theoretical underpinning.

### A. Current Assessment Practices

Translation assessment, as discussed earlier, is not a novel subject of study, of which the previous findings were worth comparing and discussing. In this research, several interesting findings unfold in relation to the current teaching and assessment practice and the lecturer expectation about the future scoring rubric. The findings were then compared against the previous research findings and the theoretical underpinning.

First, the translation was taught and tested formatively in a form of translation exercises, which were given at each meeting. It was intended to provide the students with a lot of translation exposures so they can improve their translating skills. This practice is in line with what Nord (2009), as cited by Bogotá & Marín (2013), suggested that "the translation teaching should be similar to the real practice of translation." It means that the teaching should reflect what the real translators do. Thus, their teaching practice has conformed the ideal translation teaching learning practice. In addition, this is also in line with an authenticity principle, as identified by Brown (2004), that "Tasks represent, or closely approximate, real-world tasks." Thus, the practice set out by the lecturers and the head of the department was an ideal one.

Second, a scoring rubric was currently employed by all the lecturers, although was of a different type and a different frequency of use. This difference is inevitable and natural considering the different educational background, translation teaching experience, teaching style, etc. Furthermore, this department is relatively new. In regard to different assessment practice, McAlester (2000) states that "in actual fact, we find that methods vary considerably...even between colleagues in the same department." To standardize the assessment, there should be a common thread to tie all the differences. In addition, this change should also derive from within, from critical reflection on existing practices (Aho:1997), as cited by Garant (2009).

Third, one lecturer raised the possibility of using scoring data as research data in addition to diagnostic, formative, summative (Martínez Melis & Hurtado Albir, 2001), and evaluative (Black & Wiliam, 2018) purposes. It means when the scoring was completed, the scoring results could be converted into research-appropriate data and then analyzed and converted scientifically into a research article.

According Hariyanto (2016), it is possible that the scoring rubric which was originally designed for teaching is also used for research as long as it is intended for research on translation teaching. It is because the underlying principles and purposes are different among the three fields. However, the findings in this research are contrary to this statement. The translation industry association scoring rubric was favored by the lecturers to base the scoring rubric designing on.

Fourth, the assessment criteria were also be communicated to the students. Therefore, the clear and simple information would be beneficial for the students (Brown, 2004: 257). Therefore, Brown (2004) further argues that communicating the assessment criteria is important so that the students receive maximum benefit, but it is the most complex task. Norton (2007) as cited by Hidayat (2013) state that problems are encountered by the students in dealing with assessment criteria. It means that the students found it difficult to understand the assessment criteria. This may de-motivate the students, if it is not addressed appropriately.

Fifth, several weakness were identified by the lecturers in regards to the current scoring rubric. There were at least five weaknesses, such as (1) not straightforward, (2) not precise, (3) not quite thorough, (4) not standardized, and (5) time it took to make and prepare the good scoring rubric. The weaknesses derived from the different scoring rubric the lecturers employed. The four shortcomings derived from the internal characteristics of the scoring rubric, while the other one came from external factor, the designing process. These are in line with the previous findings that practicality was chosen over the other assessment characteristics. This is different from the mainstream opinions which situate practicality as the last characteristic, such the article by Mobaraki & Aminzadeh (2012), Garant (2009), etc. These feedbacks should then be considered while designing a new scoring rubric.

Finally, it was clear from this discussion that the institution has followed an ideal and an appropriate translation education track. However, to improve the current practice, a new scoring rubric should be designed considering the feedbacks and the inputs from the current scoring rubric use. It is important for the students as well as the lecturers to have a single agreed scoring rubric so they can monitor the progress and performance in a clear and transparent manner.

**B. Designing of a Scoring Rubric**

We understand now that a new scoring rubric is required by the lecturers to incorporate in their assessment practice. It is then important to understand the lecturers' need and expectation. When the designing is based on the lecturers' real need and expectation, the standardization of the scoring rubric will not meet their resistance. In this research, those data unfolds and presents some interesting information.

First, an absolute agreement was reached among the lecturers in relation to the approach to design a customized scoring rubric. A top-down approach was selected as a method, which enables a future designer to pick one single agreed scoring rubric. This chosen rubric would then be adapted to the lecturer need and expectation. Unlike what Brookhart & Nitko (2015) states that the top-down approach starts from a conceptual framework, the lecturers stated that they wanted to adapt the existing scoring rubric instead. It may be due to the fact the institution prepare the students to work in the industry, so the common scoring rubric in the industry became their choice of rubric. Thus, this choice is more toward a pragmatic decision.

Second, the same absolute agreement was reached about a scoring rubric type which should be designed and developed in this institution, i.e. analytical scoring type. This finding is in line with Khanmohammad & Osanloo's (2009) conclusion that the translation lecturers in Iranian universities advocated the use of an objective analytical assessment, which both later translated into a specific scoring rubric. This is based on the fact that an analytical scoring rubric provides a specific and useful feedback to the students. It is different from the conclusion drawn by Garant (2009) that the assessment is now more toward a holistic approach. One possible explanation is that Garant (2009) carried the research in a European university where the language pair is somewhat approximate.

Third, a common thread that the lecturer expected was a practicality aspect. Practicality here refers to four considerations as suggested by Brown (2004), i.e. not too expensive, within appropriate time constraints, relatively easy to administer, and a scoring/evaluation procedure that is specific and time-efficient. In the institution itself, the number of the students in a class is relatively large, i.e. 25 students/class in average. They would be given a translation task each meeting, which lasted around 3 hours and two meetings every week. Thus, it would be discouraging for the lecturers when the rubric was not practical. This finding is somewhat different from what Hariyanto (2016) suggest that practicality is not main consideration in translating teaching assessment characteristics. In this research, it was found out that practicality should be the main concern when a new scoring rubric is designed.

Fourth, several considerations need to be taken into account when designing a customized new scoring rubric. The thoroughness level should be quite detailed, but not too detailed. In the operational level, they proposed four factors to consider, (1) a thorough criteria; (2) clear descriptions on which competence/sub-competence is evaluated; (3) clear and unbiased score ranges; and (4) appropriate and proportional weighing. It seems that the lecturer suggestions tend toward an industry oriented scoring rubric, as identified by Hariyanto (2016). According to Hariyanto (2016), in a scoring rubric for translation industry, there should be severity of error, categorization of error, and error weighting. However, he does not mention about the guideline and description, which typically consists of how to use rubric, error definition, error description, and other important information.

Finally, those four factors need to be taken into account when designing an ideal scoring rubric in this institution. The teaching learning process enables the use of a scoring rubric to score the student translation. The lecturers also advocate the use of objective assessments in a form of a scoring rubric with several notes regarding the weakness of the current scoring rubric and their expectation. The design itself should also consider the students as they are the center of the teaching and assessment.

## CONCLUSION AND SUGGESTION

### Conclusion

Based on the finding and discussion about the designing of a scoring rubric to assess the student translation at D3 English Department at Politeknik Negeri Malang, there were two conclusions drawn.

First, the assessment practice in this vocational higher education institution has followed an ideal assessment process, where assessment criteria were also shared with the students, but the scoring rubric type and the frequency of use still varied from one lecturer to the other. In addition, the scoring rubric use was meant for formative, summative purpose, evaluative, and research purpose. Research purpose is a new finding which has not been described before. However, there were several shortcomings in relation the current scoring rubric, which therefore requires a designing of a new scoring rubric.

Second, the designing was expected to be based on the adapting or top-down approach with a practical but detailed analytical scoring rubric as a choice. It can concluded that their expectation indicate the tendency toward a translation industry oriented scoring rubric. It can be seen from the format they expected, the components to incorporate, and the criteria they set. In short, the resulting scoring rubric is expected to be practical, valid, and reliable.

### Suggestion

In regards with the findings of the present study, a number of suggestions are proposed.

A. It is suggested that the head of the department invite all stakeholders to discuss about the weaknesses of the current rubric and its application in the classroom activities. By so doing, all the lecturers are expected to have a collective awareness about the weakness and solve the problems

B. It is suggested that the lecturers ask the institution to discuss a detailed plan to develop or adapt a scoring rubric for the maximum benefits of the students. A simple and agreed scoring rubric will let the students know and identify their weakness. Based on this awareness, the students will improve themselves.

C. It is suggested that the future researchers follows up this research by designing a new scoring rubric using a research and development method or education design research method by involving all the stakeholders and considering the actual daily problems and challenges encountered by all the lecturers. Therefore, the resulting scoring rubric will be grounded and received by the lecturers and students at Politeknik Negeri Malang.

## BIBLIOGRAPHY


Angelelli, C. V. (2009). Using a rubric to assess translation ability: Defining the construct. In C. V Angelelli & H. E. Jacobson (Eds.), *Testing and Assessment in Translation and Interpreting Studies*. Philadelpia: John Benjamins Publishing Company.

Ary, D., Jacobs, L. C., Razavieh, A., & Sorensen, C. K. (2010). *Introduction to Research in Education*. *Educational and Child Psychology*. Canada: Wadsworth, Cengage Learning ALL. https://doi.org/10.1002/ace

Barghout, M. A.-M. (1990). *Translation Quality Assessment: An Application of a Rhetorical Model*. University of Salford. Retrieved from http://usir.salford.ac.uk/2083/1/D091371.pdf

Benhaddou, M. (1991). *Translation Quality Assessment: A Situational/Textual Model for the Evaluation of Arabic/English Translations*. University of Salford. Retrieved from http://usir.salford.ac.uk/2082/1/D097677.pdf

Black, P., & Wiliam, D. (2018). Classroom assessment and pedagogy. *Assessment in Education: Principles, Policy and Practice*, 1–25. https://doi.org/10.1080/0969594X.2018.1441807

Bogotá, C. N., & Marín, P. (2013). Identifying Translation Teaching Strategies : An Exploratory Study. *International Journal of Humanities and Social Science*, *3*(21), 71–78. Retrieved from http://www.ijhssnet.com/journals/Vol_3_No_21_[Special_Issue_December_2013]/7.pdf

Bowen, G. A. (2009). Document analysis as a qualitative research method. *Qualitative Research Journal*, *9*(2), 27–40. https://doi.org/10.3316/QRJ0902027

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. https://doi.org/10.1017/CBO9781107415324.004

Brookhart, S. M. (2013). *How to create and use rubrics for formative assessment and grading*. Alexandria: ASCD.

Brookhart, S. M., & Nitko, A. J. (2015). *Educational Assessment of Students*. *Pearson Education*. New York: Pearson Education. Retrieved from https://www.pearsonhighered.com/assets/samplechapter/0/1/3/3/0133830268.pdf

Brown, H. D. (2004). *Language Assessment: Principles and Classroom Practices*. *Assessment in Education: Principles, Policy and Practice*. New York: Pearson Education, Inc.

Cuc, P. T. K. (2017). An Analysis of Translation Errors: A Case Study of Vietnamese EFL Students. *International Journal of English Linguistics*, *8*(1), 22. https://doi.org/10.5539/ijel.v8n1p22

Garant, M. (2009). A case for holistic translation assessment. *AFinLA-e: Soveltavan Kielitieteen Tutkimuksia*, *1*(2009), 5–17.

Gavotto-Nogales, O. I., Morales, L. D. G., & Pierra, L. I. C. (2015). Formative Assessment as an Essential Competence of University Teachers. *IOSR Journal of Research & Method in Education*, *5*(3), 44–47. https://doi.org/10.9790/7388-05344447

Ghafouripour, S., Branch, P., Eslamieh, R., & Branch, P. (2018). A Translation

Quality Assessment of Two English Translations of Rubaiyat of Omar Khayyam based on Juliane House's Model (1997). *International Journal of English Language & Translation Studies*, *06*(02), 217–226.

Hariyanto, S. (2009). *The Translation of Company Websites from English into Indonesian*. State University of Malang.

Hariyanto, S. (2016). Assessment in Translation Research , Teaching , and Industry. In *The 2016 International Translation and Interpreting Symposium* (pp. 236–344). Retrieved from https://www.researchgate.net/profile/Sugeng_Hariyanto3/publication/327344825_Assessment_in_Translation_Research_Teaching_and_Industry/links/5b894068299bf1d5a7339479/Assessment-in-Translation-Research-Teaching-and-Industry.pdf

Hidayat, M. T. (2013). Self - , Peer - and Teacher - assessment in Translation Course. Retrieved from http://file.upi.edu/Direktori/FPBS/JUR._PEND._BAHASA_INGGRIS/196706091994031-DIDI_SUKYADI/SELF, PEER AND TEACHER ASSESSMENT IN TRANSLATION COURSE.pdf

House, J. (2015). *Translation Quality Assessment*. (Intergovernmental Panel on Climate Change, Ed.) (First Edit). New York: Routledge. https://doi.org/10.1017/CBO9781107415324.004

Jabbarifar, T. (2009). The Importance of Classroom Assessment and Evaluation in Educational System. *Proceedings of the 2nd International Conference of Teaching and Learning (ICTL 2009)*, (Ictl), 1–9.

Khanmohammad, H., & Osanloo, M. (2009). Moving toward Objective Scoring: A Rubric for Translation Assessment. *Journal of English Language Studies*, *1*(1), 131–153. https://doi.org/10.1111/j.1600-079X.2007.00473.x

Leavy, P. (2017). *Research Design: Quantitative, Qualitative, Mixed Methods, Arts-Based, and Community-Based Participatory Research Approaches*. New York: The Guilford Press.

Martínez Melis, N., & Hurtado Albir, A. (2001). Assessment In Translation Studies: Research Needs. *Meta: Journal Des Traducteurs*, *46*(2), 272. https://doi.org/10.7202/003624ar

McAlester, G. (2000). Developing Translation Competence: Introduction. In C. Schäffner & B. Adab (Eds.), *The evaluation of translation into a foreign language* (p. vii). Amsterdam: Benjamins Translation Library. https://doi.org/10.1075/btl.38.01sch

Medadian, G., & Mahabadi, D. N. (2015). A Summative Translation Quality Assessment Model for Undergraduate Student Translations: Objectivity Versus Manageability. *Studies About Languages*, *0*(26), 40–54. https://doi.org/10.5755/j01.sal.0.26.12421

Melis, N. M., & Albir, A. H. (2001). Assessment In Translation Studies: Research Needs. *Meta: Journal Des Traducteurs*, *46*(2), 272. https://doi.org/10.7202/003624ar

Mobaraki, M., & Aminzadeh, S. (2012). A Study on Different Translation Evaluation Strategies to Introduce an Eclectic Method. *International Journal of English Linguistics*, *2*(6), 63–70. https://doi.org/10.5539/ijel.v2n6p63

Popham, W. J. (1997). What's Wrong and What's Right with Rubrics. *Educational*

*Leadership*. Retrieved from http://skidmore.edu/assessment/handbook/Popham_1997_Whats-Wrong_and-Whats-Right_With-Rubrics.pdf

Richards, J. C., & Schmidt, R. (2002). *Longman Dictionary of Language Teaching and Applied Linguistics*. London: Pearson Education Limited. https://doi.org/10.4324/9781315833835

Shiyab, S. M. (2013). Translation Quality Assessment: A Perspective on Pedagogy. *Arab World English Journal*, *2*(2), 42–50. Retrieved from www.awej.org

Stevens, D., & Lewvi, A. (2005). *Introduction to Rubrics: An Assessment Tool to Save Grading Time, Convey Effective Feedback and Promote Student Learning*. *Journal of College Student Development*. https://doi.org/10.1111/j.1541-4329.2006.00011.x

Torrance, H., & Pryor, J. (2001). Developing formative assessment in the classroom: Using action research to explore and modify theory. *British Educational Research Journal*. https://doi.org/10.1080/01411920120095780

Waddington, C. (2003). A positive approach to the assessment of translation errors. *Granada: AIETI*, (2), 409–426.

William, D. (2013). Assessment : The Bridge between Teaching and Learning. *Voices from the Middle*, *21*(2), 40. Retrieved from https://www.ncte.org/library/NCTEFiles/Resources/Journals/VM/0212-dec2013/VM0212Assessment.pdf

Zehnalová, J. (2013). Tradition and Trends in Translation Quality Assessment. In *Tradition and Trends in Trans-Language Communication* (pp. 41–58). Olomouc: Univerzita Palackého.